

Figure 1

Heteractis crispa chromoprotein **wild type (base isoform)**

```

      10      20      30      40      50      60
5' ACCATTTGCTTTGGTTCCTTGGCAAACGAAAGTTTAGAACGAAACTGACCCAAATTACA
      70      80      90     100     110     120
   TCTTCCTCCTGGATCCTTACCATGGCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATG
               M A G L L K E S M R I K M
      130     140     150     160     170     180
   TACATGGAAGGCACGGTTAATGGCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAAC
   Y M E G T V N G H Y F K C E G E G D G N
      190     200     210     220     230     240
   CCATTACAGGTACGCAGAGCATGAGGATTCATGTCACCGAAGGGGCTCCATTACCATT
   P F T G T Q S M R I H V T E G A P L P F
      250     260     270     280     290     300
   GCCTTCGACATTTTGGCACCGTGTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACG
   A F D I L A P C C E Y G S R T F V H H T
      310     320     330     340     350     360
   GCAGAGATTCCCGATTTCTTCAAGCAGTCTTTCCTGAAGGCTTTACTTGGGAAAGAACC
   A E I P D F F K Q S F P E G F T W E R T
      370     380     390     400     410     420
   ACAACCTATGAAGATGGAGGCATTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAAC
   T T Y E D G G I L T A H Q D T S L E G N
      430     440     450     460     470     480
   TGCCTTATATAACAAGGTGAAAGTCCTTGGTACCAATTTTCCTGCTGATGGCCCCGTGATG
   C L I Y K V K V L G T N F P A D G P V M
      490     500     510     520     530     540
   AAGAACAAATCAGGAGGATGGGAGCCATGCACTGAGGTGGTTTATCCAGAGAATGGTGTC
   K N K S G G W E P C T E V V Y P E N G V
      550     560     570     580     590     600
   CTGTGTGGACGTAATGTGATGGCCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCTC
   L C G R N V M A L K V G D R R L I C H L
      610     620     630     640     650     660
   TATACTTCTTACAGGTCCAAGAAAGCAGTCCGTGCCTTGACAATGCCAGGATTTTCATTTT
   Y T S Y R S K K A V R A L T M P G F H F
      670     680     690     700     710     720
   ACAGACATCCGCCTTCAGATGCCGAGGAAAAAGAAAGACGAGTACTTTGAACTGTACGAA
   T D I R L Q M P R K K K D E Y F E L Y E
      730     740     750     760     770     780
   GCATCTGTGGCTAGGTACAGTGATCTTCCTGAAAAAGCAAATTGATTGTTCCCAGTGACA
   A S V A R Y S D L P E K A N *
      790     800     810     820     830     840
   CCAGACTGCTGTCAGCTTTTGGTTAAAGCCCGAAAGACAAAAGGACATTTGTAGTTTAGT
      850     860     870     880     890     900
   TTATATTTCCCTTTTCATTTGTGAATCAACATTGTACTCTCTGTAAACCTTTAAATGCTC
      910
   CATTAAACCT 3' (SEQ ID NOS: 01 & 02)

```

099697-104204

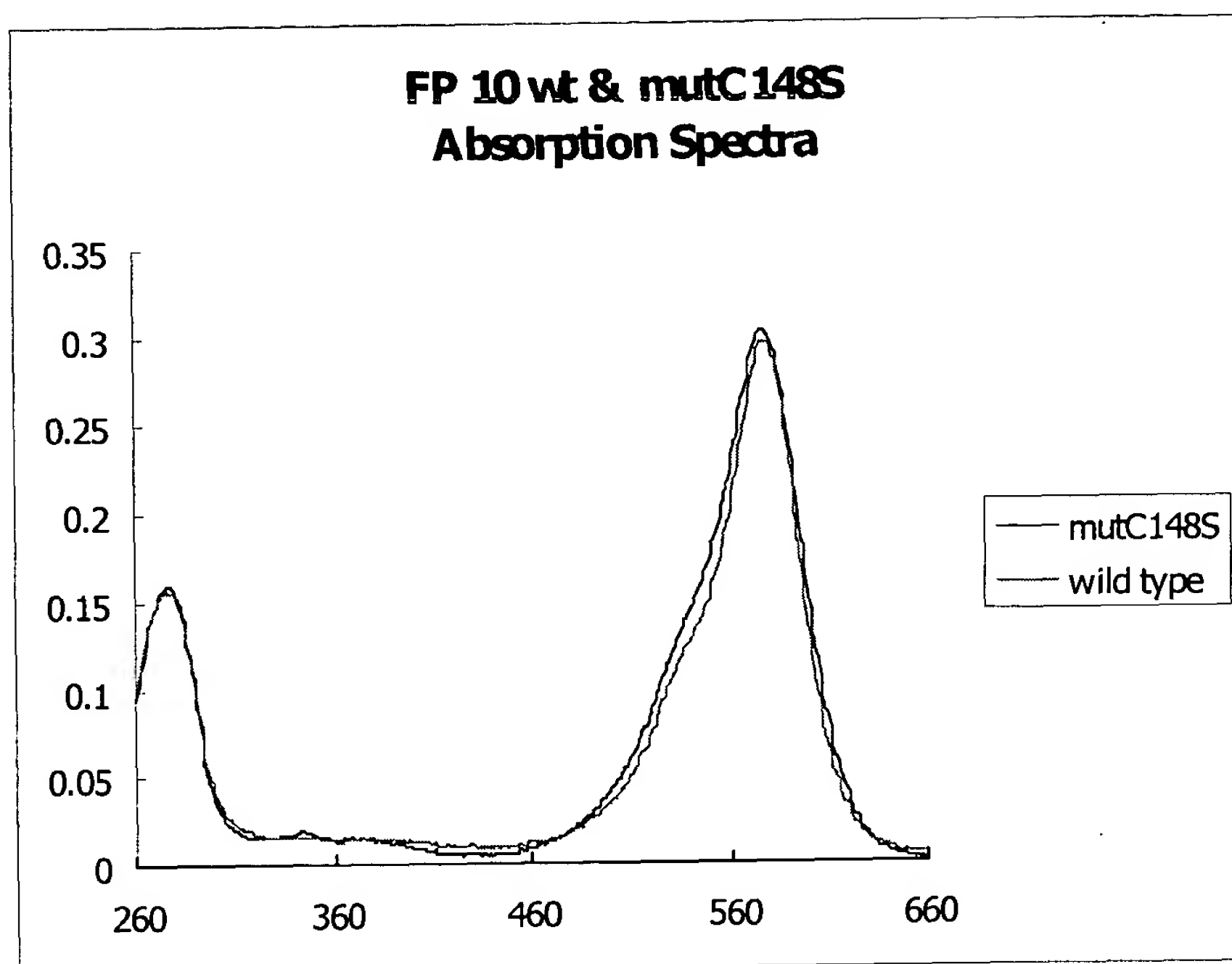
Figure 2

Heteractis crispa chromoprotein wild type (second isoform)

10 20 30 40 50 60
5' ACCATTTGCTTTGGTTCCTTGGCAAACGAAAGTTTAGACGAAAAGTACCCAAATTACAT
70 80 90 100 110 120
CCTCCTGATCCTTACCATGGCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACAT
M A G L L K E S M R I K M Y M
130 140 150 160 170 180
GGAAGGCACGGTTAATGGCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATT
E G T V N G H Y F K C E G E G D G N P F
190 200 210 220 230 240
TACAGGTACGCAGAGCATGAGGATTCATGTCACCGAAGGGGCTCCATTACCATTTGCCTT
T G T Q S M R I H V T E G A P L P F A F
250 260 270 280 290 300
CGACATTTTGGCACCGTGTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGA
D I L A P C C E Y G S R T F V H H T A E
310 320 330 340 350 360
GATTCCCGATTTCTTCAAGCAGTCTTTCCCTGAAGGCTTTACTTGGGAAAGAACCACAAC
I P D F F K Q S F P E G F T W E R T T T
370 380 390 400 410 420
CTATGAAGATGGAGGCATTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAACTGCCT
Y E D G G I L T A H Q D T S L E G N C L
430 440 450 460 470 480
TATATACAAGGTGAAAGTCCTTGGTACCAATTTTCCTGCTGATGGCCCCGTGATGAAGAA
I Y K V K V L G T N F P A D G P V M K N
490 500 510 520 530 540
CAAATCAGAAGGATGGGAGCCATGCACTGAGGTGGTTTATCCAGATAATGGTGTCTCTGTG
K S E G W E P C T E V V Y P D N G V L C
550 560 570 580 590 600
TGGACGTAATGTGATGGCCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCTCTATAC
G R N V M A L K V G D R R L I C H L Y T
610 620 630 640 650 660
TTCTTACAGGTCCAAGAAAGCAGTCCGTGCCTTGACAATGCCAGGATTTTCAATTTACAGA
S Y R S K K A V R A L T M P G F H F T D
670 680 690 700 710 720
CATCCGCCTTCAGATGCCGAGGAAAAAGAAAGACGAGTACTTTGAACTGTACGAAGCATC
I R L Q M P R K K K D E Y F E L Y E A S
730 740 750 760 770 780
TGTGGCTAGGTACAGTGATCTTCTGAAAAAGCAAATTGATTGTTCCAGTGACACCAGA
V A R Y S D L P E K A N *
790 800 810 820 830 840
CTGCTGTCAGCTTTTGGTTAAAGCCCGAAAGACAAAAGGACATTTGTAGTTTTAGTTTAT
850 860 870 880 890 900
ATTTTCCCTTTCATTTTGTGAATCAACATTGTACTCTCTGTAAACCTTTAAAATGCTCCA
TTAAACCT 3' (SEQ ID NOS: 03 & 04)

09976673-101201

Figure 3



109763-10404
F02T01 E294560

Figure 4

Heteractis crispa fluorescent protein mutant **C148S**

C148S according to GFP numbering
C143S according to self-numbering.

ATGGCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATG
M A G L L K E S M R I K M

TACATGGAAGGCACGGTTAATGGCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAAC
Y M E G T V N G H Y F K C E G E G D G N

CCATTTACAGGTACGCAGAGCATGAGGATTCATGTCACCGAAGGGGCTCCATTACCATTT
P F T G T Q S M R I H V T E G A P L P F

GCCTTCGACATTTTGGCACCGTGTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACG
A F D I L A P C C E Y G S R T F V H H T

GCAGAGATTCCCGATTTCTTCAAGCAGTCTTCCCTGAAGGCTTTACTTGGGAAAGAACC
A E I P D F F K Q S F P E G F T W E R T

ACAACCTATGAAGATGGAGGCATTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAAC
T T Y E D G G I L T A H Q D T S L E G N

TGCCTTATATACAAGGTGAAAGTCCTTGGTACCAATTTTCCTGCTGATGGCCCCGTGATG
C L I Y K V K V L G T N F P A D G P V M

AAGAACAAATCAGGAGGATGGGAGCCAAGCACTGAGGTGGTTTATCCAGAGAATGGTGTC
K N K S G G W E P S T E V V Y P E N G V

CTGTGTGGACGTAATGTGATGGCCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCTC
L C G R N V M A L K V G D R R L I C H L

TATACTTCTTACAGGTCCAAGAAAGCAGTCCGTGCCTTGACAATGCCAGGATTTTCATTTT
Y T S Y R S K K A V R A L T M P G F H F

ACAGACATCCGCCTTCAGATGCCGAGGAAAAAGAAAGACGAGTACTTTGAACTGTACGAA
T D I R L Q M P R K K K D E Y F E L Y E

GCATCTGTGGCTAGGTACAGTGATCTTCCTGAAAAAGCAAATTGA
A S V A R Y S D L P E K A N *

(SEQ ID NOS: 05 & 06)

0997663-101301

Figure 5A

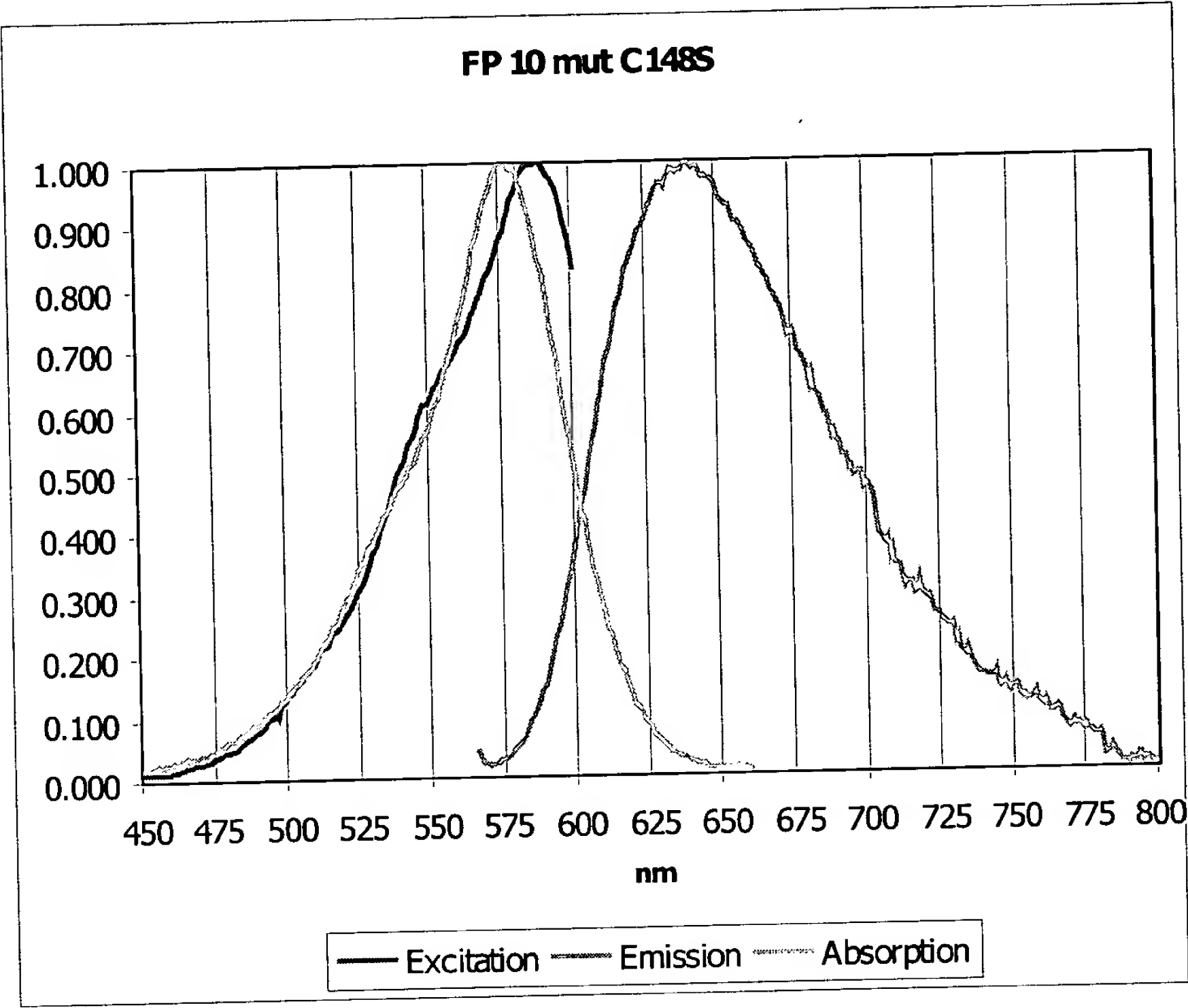


Figure 5B

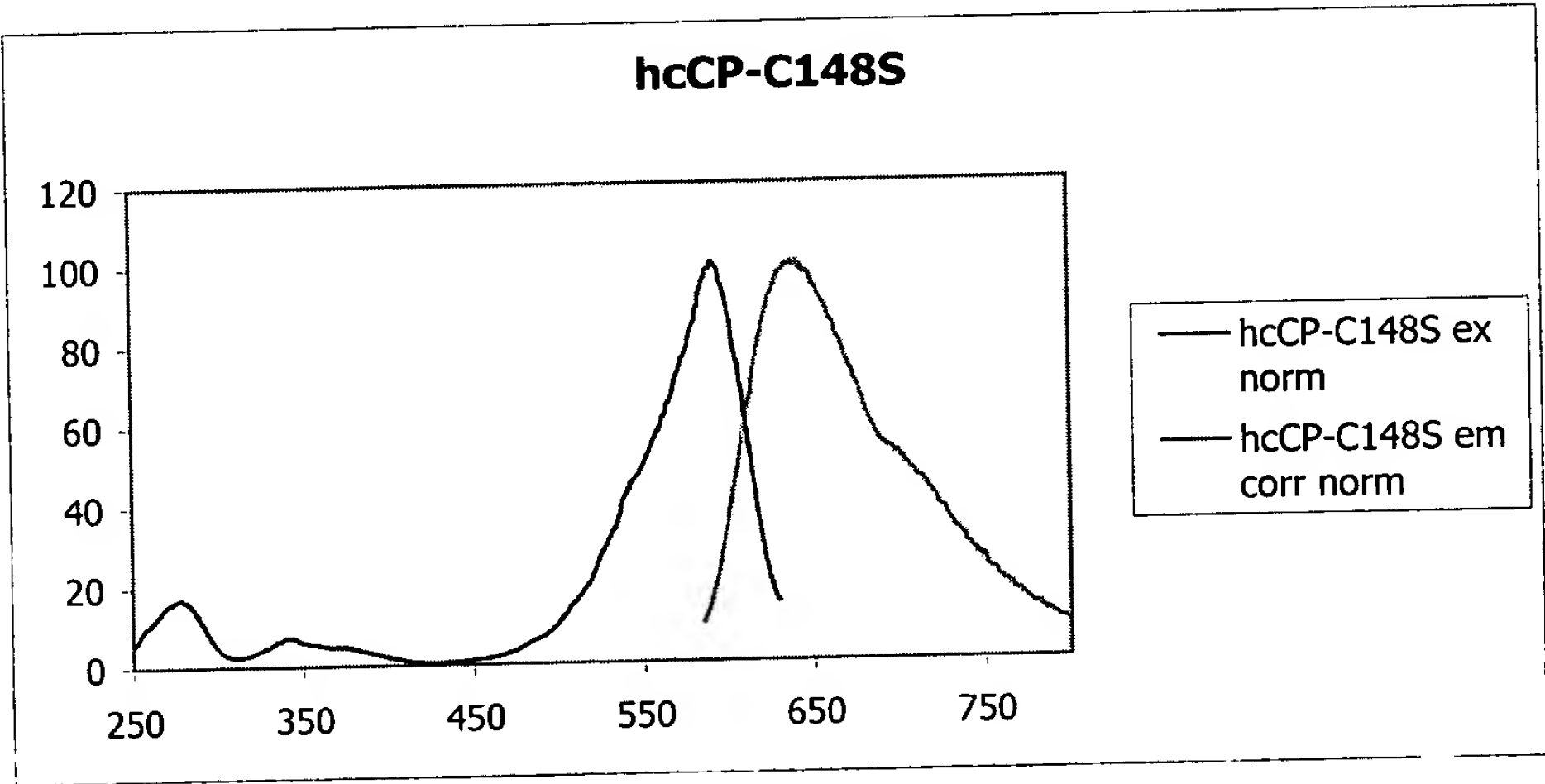


Figure 6

Heteractis crispa fluorescent protein mutant 44-9

point mutations: A5S,T39A,C148S,L181H,P208L,K211E according to GFP numbering

A2S,T36A,C143S,L173H,P201L,K204E according to self-numbering.

```

      80      90      100      110      120
      TCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACAT
      S G L L K E S M R I K M Y M

      130      140      150      160      170      180
GGAAGGCACGGTTAATGGCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATT
E G T V N G H Y F K C E G E G D G N P F

      190      200      210      220      230      240
TGCAGGTACGCAGAGCATGAGGATTCATGTCACCGAAGGGGCTCCATTACCATTTCGCTT
A G T Q S M R I H V T E G A P L P F A F

      250      260      270      280      290      300
CGACATTTTGGCACC GTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATAACGGCAGA
D I L A P C C E Y G S R T F V H H T A E

      310      320      330      340      350      360
GATTCCCGATTTCTTCAAGCAGTCTTTCCTGAAGGCTTTACTTGGGAAAGAACCACAAC
I P D F F K Q S F P E G F T W E R T T T

      370      380      390      400      410      420
CTATGAAGATGGAGGCATTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAACTGCCT
Y E D G G I L T A H Q D T S L E G N C L

      430      440      450      460      470      480
TATATACAAGGTGAAAGTCCTTGGTACCAATTTTCCTGCTGATGGCCCCGTGATGAAGAA
I Y K V K V L G T N F P A D G P V M K N

      490      500      510      520      530      540
CAAATCAGGAGGATGGGAGCCAAGCACTGAGGTGGTTTATCCAGAGAATGGTGTCTCTGTG
K S G G W E P S T E V V Y P E N G V L C

      550      560      570      580      590      600
TGGACGTAATGTGATGGCCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCACTATAC
G R N V M A L K V G D R R L I C H H Y T

      610      620      630      640      650      660
TTCTTACAGGTCCAAGAAAGCAGTCCGTGCCTTGACAATGCCAGGATTTCATTTTACAGA
S Y R S K K A V R A L T M P G F H F T D

      670      680      690      700      710      720
CATCCGCCTTCAGATGCTGAGGAAAGAGAAAGACGAGTACTTTGAACTGTACGAAGCATC
I R L Q M L R K E K D E Y F E L Y E A S

      730      740      750      760
TGTGGCTAGGTACAGTGATCTTCCTGAAAAAGCAAATTGA
V A R Y S D L P E K A N * (SEQ ID NOs: 07 & 08)
```

099697401

Figure 7A

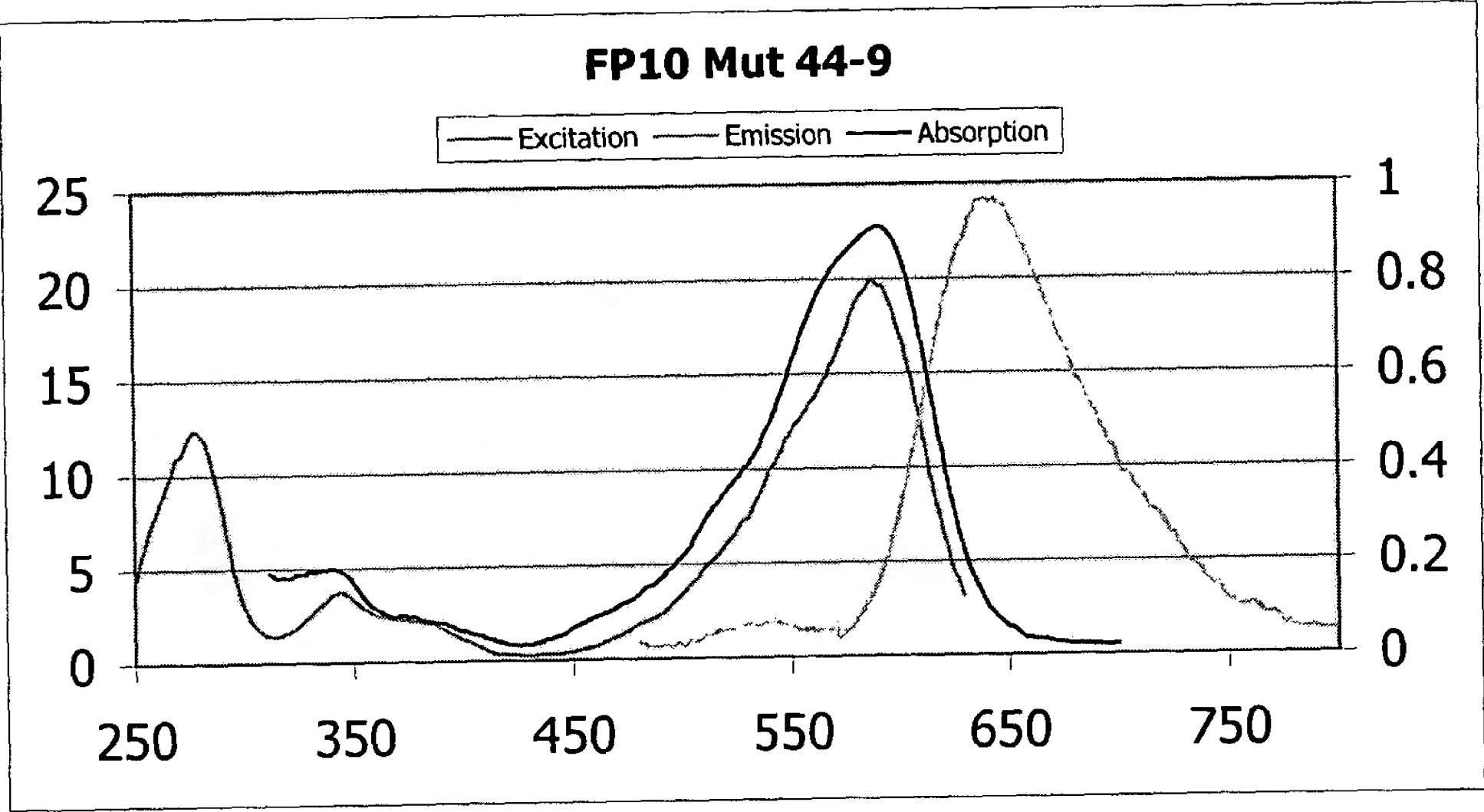


Figure 7B

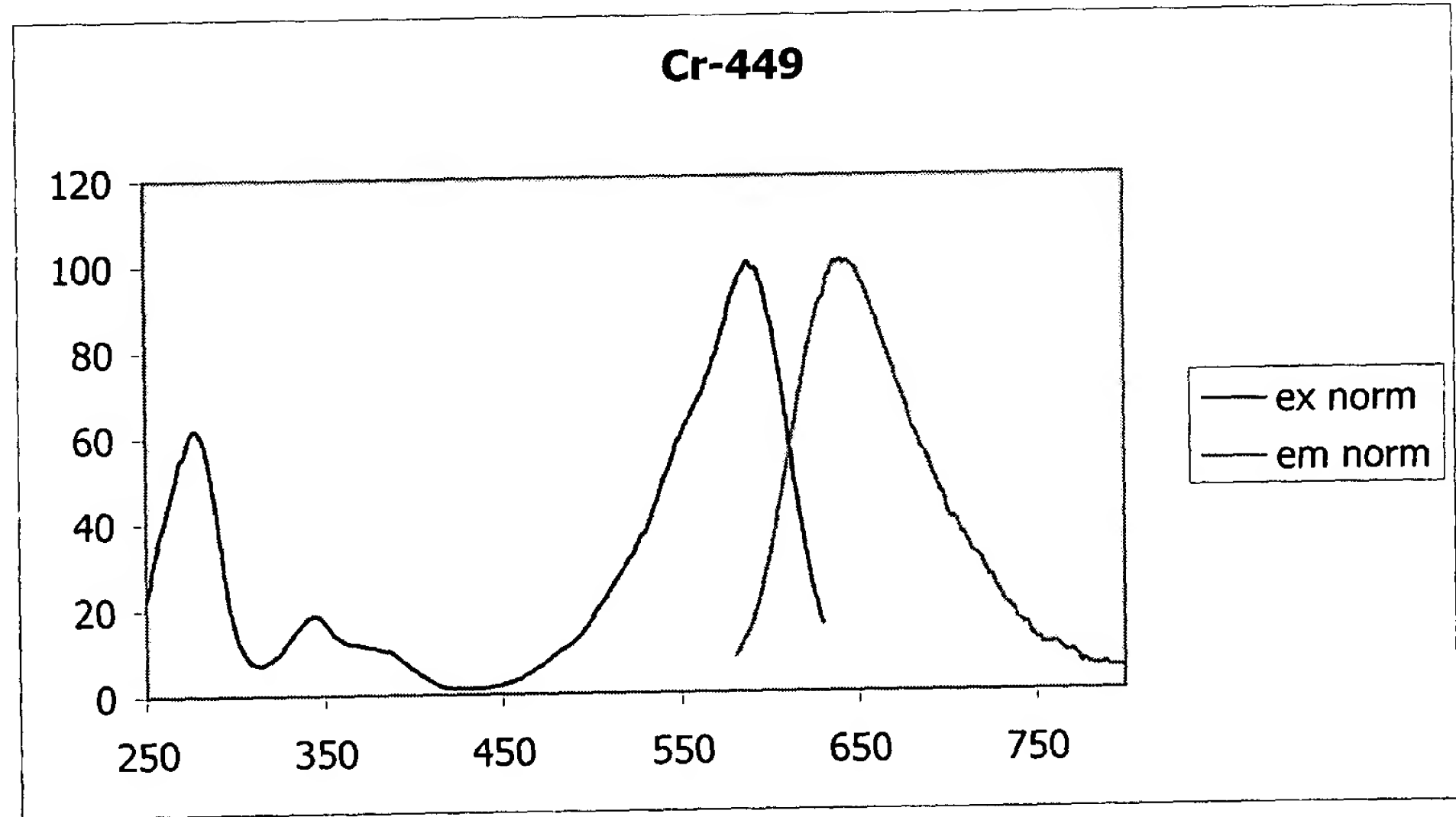


Figure 8

**Crispa 44-6 mutant possesses six amino acid substitutions vs. wild type:
A2S,T36A,A65E,C143S,L173H,P201L.**

TCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACAT
S G L L K E S M R I K M Y M

GGAAGGCACGGTTAATGGCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATT
E G T V N G H Y F K C E G E G D G N P F

TGCAGGTACGCAGAGCATGAGGATTCATGTCACCGAAGGGGCTCCATTACCATTTCGCTT
A G T Q S M R I H V T E G A P L P F A F

CGACATTTTGGCACCGTGTGTGCGTACGGCAGCAGGACCTTTGTCCACCATAACGGCAGA
D I L A P C C A Y G S R T F V H H T A E

GATTCCCGATTTCCTTCAAGCAGTCTTTCCTGAAGGCTTTACTTGGGAAAGAACCACAAC
I P D F F K Q S F P E G F T W E R T T T

CTATGAAGATGGAGGCATTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAAGTGCCT
Y E D G G I L T A H Q D T S L E G N C L

←
TATATACAAGGTGAAAGTCCTTGGTACCAATTTTCCTGCTGATGGCCCCGTGATGAAGAA
I Y K V K V L G T N F P A D G P V M K N

CAAATCAGGAGGATGGGAGCCAAGCACTGAGGTGGTTTATCCAGAGAATGGTGTCTCTGTG
K S G G W E P S T E V V Y P E N G V L C

TGGACGTAATGTGATGGCCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCACTATAC
G R N V M A L K V G D R R L I C H H Y T

TTCTTACAGGTCCAAGAAAGCAGTCCGTGCCTTGACAATGCCAGGATTTTCATTTTACAGA
S Y R S K K A V R A L T M P G F H F T D

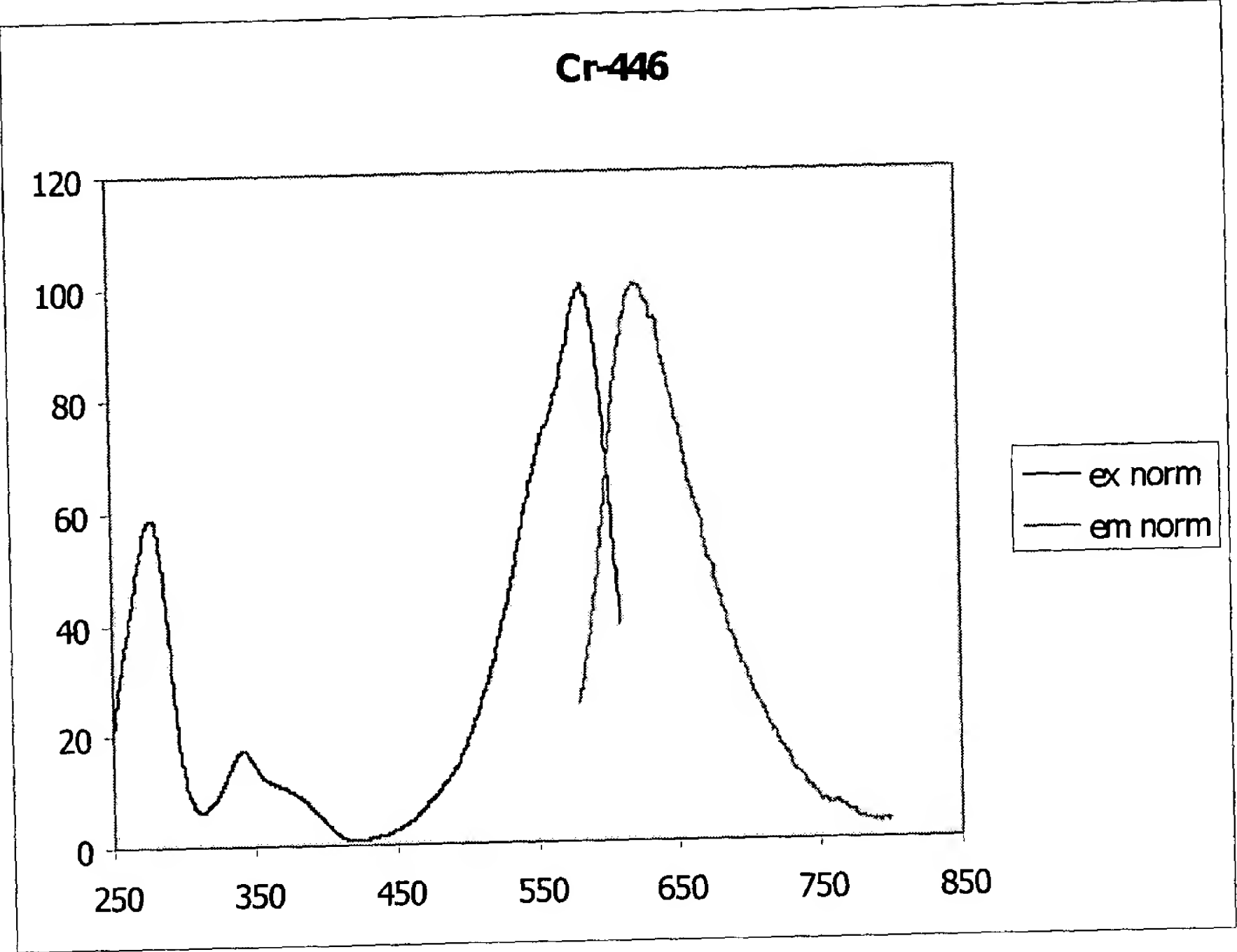
CATCCGCCTTCAGATGCTGAGGAAAGAGAAAGACGAGTACTTTGAACTGTACGAAGCATC
I R L Q M L R K E K D E Y F E L Y E A S

TGTGGCTAGGTACAGTGATCTTCCTGAAAAAGCAAATTGA
V A R Y S D L P E K A N *

(SEQ ID NO: 09 & 10)

09976673-101204

Figure 9



09963-10404
FOR "E49260"

Figure 10

The amino acid sequence of FP10-cr1 is:

M S G L L K E S M R I K M Y M E G T V N G H Y
F K C E G E G D G N P F A G T Q S M R I H V T
E G A P L P F A F D I L A P C C E Y G S R T F
V H H T A E I P D F F K Q S F P E G F T W E R
T T T Y E D G G I L T A H Q D T S L E G N C L
I Y K V K V H G T N F P A D G P V M K N K S G
G W E P S T E V V Y P E N G V L C G R N V M A
L K V G D R H L I C H H Y T S Y R S K K A V R
A L T M P G F H F T D I R L Q M L R K K K D E
Y F E L Y E A S V A R Y S D L P E K A N (SEQ ID
NO:12)

A "humanized" nt sequence encoding the above cr-1 mutant is:

ATGGTGAGCGGCCTGCTGAAGGAGAGTATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGGCCAC
TACTTCAAGTGCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCATGAGAATCCACGTG
ACCGAGGGCGCCCCCTGCCCTTCGCCTTCGACATCCTGGCCCCCTGCTGCGAGTACGGCAGCAGGACC
TTCGTGCACCACACCGCCGAGATCCCCGACTTCTTCAAGCAGAGCTTCCCCGAGGGCTTCACCTGGGAG
AGAACCACCACCTACGAGGACGGCGGCATCCTGACCGCCCACCAGGACACCAGCCTGGAGGGCAACTGC
CTGATCTACAAGGTGAAGGTGCACGGCACCAACTTCCCCGCCGACGGCCCCGTGATGAAGAACAAGAGC
GGCGGCTGGGAGCCCAGCACCGAGGTGGTGTACCCCGAGAACGGCGTGCTGTGCGGCCGGAACGTGATG
GCCCTGAAGGTGGGCGACCGGCACCTGATCTGCCACCACTACACCAGCTACCGGAGCAAGAAGGCCGTG
CGCGCCCTGACCATGCCCCGGCTTCCACTTCACCGACATCCGGCTCCAGATGCTGCGGAAGAAGAAGGAC
GAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCCGTACAGCGACCTGCCCCGAGAAGGCCAACTGA
(SEQ ID NO:11)

Alternative cr1 amino acid sequence

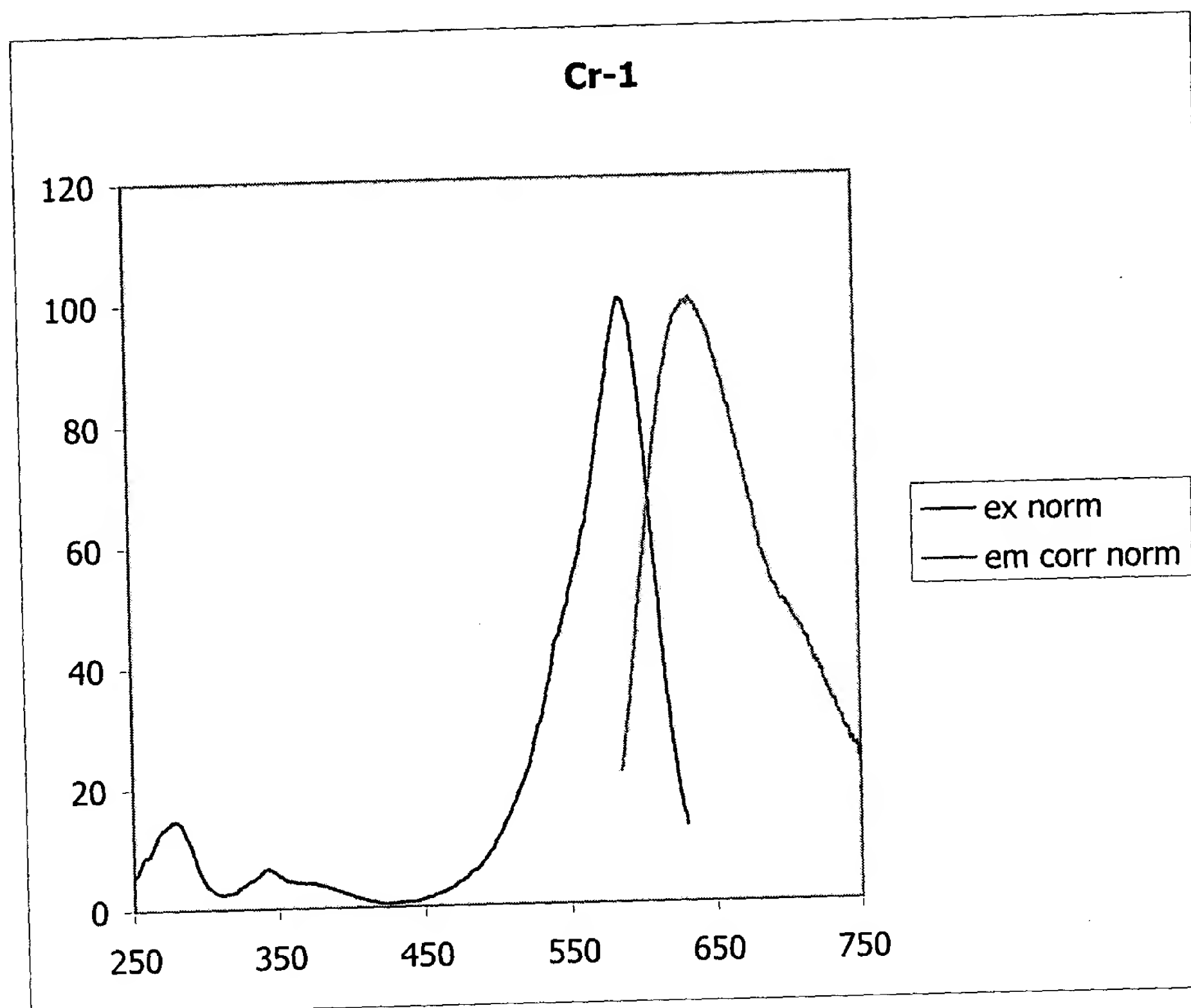
MSGLLKESMRIKMYMEGTVNGHYFKCEGEGDGNPFAGTQSMRIHVTEGAPLPFAFDILAPCCEYGSRTF
VHHTAEIPDFFKQSFPEGFTWERTTTYEDGGILTAAHQDTSLEGNCLIIYKVKVHGTNFPADGPVMKNKSG
GWEPSTEVVYPENGVLCGRNVMALKVGDRLICHHYTSYRSKKAVRALTMPGFHFDTDIRLQMLRKEKDE
YFELYEASVARYSDLPKAN* SEQ ID NO:14)

Amino acid sequence encoding above alternative sequence:

ATGGTGAGCGGCCTGCTGAAGGAGAGCATGCGCATCAAGATGTACATGGAGGGCACCGTGAACGGCCAC
TACTTCAAGTGCGAGGGCGAGGGCGACGGCAACCCCTTCGCCGGCACCCAGAGCATGCGGATCCACGTG
ACCGAGGGCGCCCCCTGCCCTTCGCCTTCGACATCCTGGCCCCCTGCTGCGAGTACGGCAGCAGGACC
TTCGTGCACCACACCGCCGAGATCCCCGACTTCTTCAAGCAGAGCTTCCCCGAGGGCTTCACCTGGGAG
AGAACCACCACCTACGAGGACGGCGGCATCCTGACCGCCCACCAGGACACCAGCCTGGAGGGCAACTGC
CTGATCTACAAGGTGAAGGTGCTGGGCACCAACTTCCCCGCCGACGGCCCCGTGATGAAGAACAAGAGC
GGCGGCTGGGAGCCCAGCACCGAGGTGGTGTACCCCGAGAACGGCGTGCTGTGCGGCCGGAACGTGATG
GCCCTGAAGGTGGGCGACCGGCGGCTGATCTGCCACCACTACACCAGCTACCGGAGCAAGAAGGCCGTG
CGGGCCCTGACCATGCCCCGGCTTCCACTTCACCGACATCCGGCTGCAGATGCTGCGGAAGGAGAAGGAC
GAGTACTTCGAGCTGTACGAGGCCAGCGTGGCCCCGTACAGCGACCTGCCCCGAGAAGGCCAACTGA
(SEQ ID NO:13)

09976673.101201

Figure 11



09976673-101204

FIGURE 12

Cr-449-tandem (4-amino acid linker between monomers is in double underline).

1	A	CCG	GTC	GCC	ACC	ATG	GTG	AGC	GGC	CTG	CTG	AAG	GAG	AGC	ATG	CGC	46
1	AgeI					M	V	S	G	L	L	K	E	S	M	R	11
47	ATC	AAG	ATG	TAC	ATG	GAG	GGC	ACC	GTG	AAC	GGC	CAC	TAC	TTC	AAG	TGC	94
12	I	K	M	Y	M	E	G	T	V	N	G	H	Y	F	K	C	27
95	GAG	GGC	GAG	GGC	GAC	GGC	AAC	CCC	TTC	GCC	GGC	ACC	CAG	AGC	ATG	CGG	142
28	E	G	E	G	D	G	N	P	F	A	G	T	Q	S	M	R	43
143	ATC	CAC	GTG	ACC	GAG	GGC	GCC	CCC	CTG	CCC	TTC	GCC	TTC	GAC	ATC	CTG	190
44	I	H	V	T	E	G	A	P	L	P	F	A	F	D	I	L	59
191	GCC	CCC	TGC	TGC	GAG	TAC	GGC	AGC	AGG	ACC	TTC	GTG	CAC	CAC	ACC	GCC	238
60	A	P	C	C	E	Y	G	S	R	T	F	V	H	H	T	A	75
239	GAG	ATC	CCC	GAC	TTC	TTC	AAG	CAG	AGC	TTC	CCC	GAG	GGC	TTC	ACC	TGG	286
76	E	I	P	D	F	F	K	Q	S	F	P	E	G	F	T	W	91
287	GAG	AGA	ACC	ACC	ACC	TAC	GAG	GAC	GGC	GGC	ATC	CTG	ACC	GCC	CAC	CAG	334
92	E	R	T	T	T	Y	E	D	G	G	I	L	T	A	H	Q	107
335	GAC	ACC	AGC	CTG	GAG	GGC	AAC	TGC	CTG	ATC	TAC	AAG	GTG	AAG	GTG	CTG	382
108	D	T	S	L	E	G	N	C	L	I	Y	K	V	K	V	L	123
383	GGC	ACC	AAC	TTC	CCC	GCC	GAC	GGC	CCC	GTG	ATG	AAG	AAC	AAG	AGC	GGC	430
124	G	T	N	F	P	A	D	G	P	V	M	K	N	K	S	G	139
431	GGC	TGG	GAG	CCC	AGC	ACC	GAG	GTG	GTG	TAC	CCC	GAG	AAC	GGC	GTG	CTG	478
140	G	W	E	P	S	T	E	V	V	Y	P	E	N	G	V	L	155
479	TGC	GGC	CGG	AAC	GTG	ATG	GCC	CTG	AAG	GTG	GGC	GAC	CGG	CGG	CTG	ATC	526
156	C	G	R	N	V	M	A	L	K	V	G	D	R	R	L	I	171
527	TGC	CAC	CAC	TAC	ACC	AGC	TAC	CGG	AGC	AAG	AAG	GCC	GTG	CGG	GCC	CTG	574
172	C	H	H	Y	T	S	Y	R	S	K	K	A	V	R	A	L	187
575	ACC	ATG	CCC	GGC	TTC	CAC	TTC	ACC	GAC	ATC	CGG	CTG	CAG	ATG	CTG	CGG	622
188	T	M	P	G	F	H	F	T	D	I	R	L	Q	M	L	R	203
623	AAG	GAG	AAG	GAC	GAG	TAC	TTC	GAG	CTG	TAC	GAG	GCC	AGC	GTG	GCC	CGG	670
204	K	E	K	D	E	Y	F	E	L	Y	E	A	S	V	A	R	219
671	TAC	AGC	GAC	CTG	CCC	GAG	AAG	GCC	AAC	<u>AGA</u>	<u>TCT</u>	<u>CCC</u>	<u>GGG</u>	ATG	GTG	AGC	718
220	Y	S	D	L	P	E	K	A	N	R	S	P	G	M	V	S	235
719	GGC	CTG	CTG	AAG	GAG	AGC	ATG	CGC	ATC	AAG	ATG	TAC	ATG	GAG	GGC	ACC	766
236	G	L	L	K	E	S	M	R	I	K	M	Y	M	E	G	T	251
767	GTG	AAC	GGC	CAC	TAC	TTC	AAG	TGC	GAG	GGC	GAG	GGC	GAC	GGC	AAC	CCC	814
252	V	N	G	H	Y	F	K	C	E	G	E	G	D	G	N	P	267
815	TTC	GCC	GGC	ACC	CAG	AGC	ATG	CGG	ATC	CAC	GTG	ACC	GAG	GGC	GCC	CCC	862
268	F	A	G	T	Q	S	M	R	I	H	V	T	E	G	A	P	283
863	CTG	CCC	TTC	GCC	TTC	GAC	ATC	CTG	GCC	CCC	TGC	TGC	GAG	TAC	GGC	AGC	910
284	L	P	F	A	F	D	I	L	A	P	C	C	E	Y	G	S	299

09076673-104204

Figure 12 (continued)

911	AGG	ACC	TTC	GTG	CAC	CAC	ACC	GCC	GAG	ATC	CCC	GAC	TTC	TTC	AAG	CAG	958
300	R	T	F	V	H	H	T	A	E	I	P	D	F	F	K	Q	315
959	AGC	TTC	CCC	GAG	GGC	TTC	ACC	TGG	GAG	AGA	ACC	ACC	ACC	TAC	GAG	GAC	1006
316	S	F	P	E	G	F	T	W	E	R	T	T	T	Y	E	D	331
1007	GGC	GGC	ATC	CTG	ACC	GCC	CAC	CAG	GAC	ACC	AGC	CTG	GAG	GGC	AAC	TGC	1054
332	G	G	I	L	T	A	H	Q	D	T	S	L	E	G	N	C	347
1055	CTG	ATC	TAC	AAG	GTG	AAG	GTG	CTG	GGC	ACC	AAC	TTC	CCC	GCC	GAC	GGC	1102
348	L	I	Y	K	V	K	V	L	G	T	N	F	P	A	D	G	363
1103	CCC	GTG	ATG	AAG	AAC	AAG	AGC	GGC	GGC	TGG	GAG	CCC	AGC	ACC	GAG	GTG	1150
364	P	V	M	K	N	K	S	G	G	W	E	P	S	T	E	V	379
1151	GTG	TAC	CCC	GAG	AAC	GGC	GTG	CTG	TGC	GGC	CGG	AAC	GTG	ATG	GCC	CTG	1198
380	V	Y	P	E	N	G	V	L	C	G	R	N	V	M	A	L	395
1199	AAG	GTG	GGC	GAC	CGG	CGG	CTG	ATC	TGC	CAC	CAC	TAC	ACC	AGC	TAC	CGG	1246
396	K	V	G	D	R	R	L	I	C	H	H	Y	T	S	Y	R	411
1247	AGC	AAG	AAG	GCC	GTG	CGG	GCC	CTG	ACC	ATG	CCC	GGC	TTC	CAC	TTC	ACC	1294
412	S	K	K	A	V	R	A	L	T	M	P	G	F	H	F	T	427
1295	GAC	ATC	CGG	CTG	CAG	ATG	CTG	CGG	AAG	GAG	AAG	GAC	GAG	TAC	TTC	GAG	1342
428	D	I	R	L	Q	M	L	R	K	E	K	D	E	Y	F	E	443
1343	CTG	TAC	GAG	GCC	AGC	GTG	GCC	CGG	TAC	AGC	GAC	CTG	CCC	GAG	AAG	GCC	1390
444	L	Y	E	A	S	V	A	R	Y	S	D	L	P	E	K	A	459
1391	AAC	TGA															
460	N	*															

(SEQ ID NOS. 15 & 16)

09976673-101201

Figure 13

Cr-449-tandem-actin (4-amino acid linker between Cr-449 monomers is noted in double underline; 4-amino acid linker between second Cr-449 and actin is noted in dashed underline).

1	A	<u>CCG</u>	GTC	GCC	ACC	ATG	GTG	AGC	GGC	CTG	CTG	AAG	GAG	AGC	ATG	CGC	46
1	AgeI					M	V	S	G	L	L	K	E	S	M	R	11
47	ATC	AAG	ATG	TAC	ATG	GAG	GGC	ACC	GTG	AAC	GGC	CAC	TAC	TTC	AAG	TGC	94
12	I	K	M	Y	M	E	G	T	V	N	G	H	Y	F	K	C	27
95	GAG	GGC	GAG	GGC	GAC	GGC	AAC	CCC	TTC	GCC	GGC	ACC	CAG	AGC	ATG	CGG	142
28	E	G	E	G	D	G	N	P	F	A	G	T	Q	S	M	R	43
143	ATC	CAC	GTG	ACC	GAG	GGC	GCC	CCC	CTG	CCC	TTC	GCC	TTC	GAC	ATC	CTG	190
44	I	H	V	T	E	G	A	P	L	P	F	A	F	D	I	L	59
191	GCC	CCC	TGC	TGC	GAG	TAC	GGC	AGC	AGG	ACC	TTC	GTG	CAC	CAC	ACC	GCC	238
60	A	P	C	C	E	Y	G	S	R	T	F	V	H	H	T	A	75
239	GAG	ATC	CCC	GAC	TTC	TTC	AAG	CAG	AGC	TTC	CCC	GAG	GGC	TTC	ACC	TGG	286
76	E	I	P	D	F	F	K	Q	S	F	P	E	G	F	T	W	91
287	GAG	AGA	ACC	ACC	ACC	TAC	GAG	GAC	GGC	GGC	ATC	CTG	ACC	GCC	CAC	CAG	334
92	E	R	T	T	T	Y	E	D	G	G	I	L	T	A	H	Q	107
335	GAC	ACC	AGC	CTG	GAG	GGC	AAC	TGC	CTG	ATC	TAC	AAG	GTG	AAG	GTG	CTG	382
108	D	T	S	L	E	G	N	C	L	I	Y	K	V	K	V	L	123
383	GGC	ACC	AAC	TTC	CCC	GCC	GAC	GGC	CCC	GTG	ATG	AAG	AAC	AAG	AGC	GGC	430
124	G	T	N	F	P	A	D	G	P	V	M	K	N	K	S	G	139
431	GGC	TGG	GAG	CCC	AGC	ACC	GAG	GTG	GTG	TAC	CCC	GAG	AAC	GGC	GTG	CTG	478
140	G	W	E	P	S	T	E	V	V	Y	P	E	N	G	V	L	155
479	TGC	GGC	CGG	AAC	GTG	ATG	GCC	CTG	AAG	GTG	GGC	GAC	CGG	CGG	CTG	ATC	526
156	C	G	R	N	V	M	A	L	K	V	G	D	R	R	L	I	171
527	TGC	CAC	CAC	TAC	ACC	AGC	TAC	CGG	AGC	AAG	AAG	GCC	GTG	CGG	GCC	CTG	574
172	C	H	H	Y	T	S	Y	R	S	K	K	A	V	R	A	L	187
575	ACC	ATG	CCC	GGC	TTC	CAC	TTC	ACC	GAC	ATC	CGG	CTG	CAG	ATG	CTG	CGG	622
188	T	M	P	G	F	H	F	T	D	I	R	L	Q	M	L	R	203
623	AAG	GAG	AAG	GAC	GAG	TAC	TTC	GAG	CTG	TAC	GAG	GCC	AGC	GTG	GCC	CGG	670
204	K	E	K	D	E	Y	F	E	L	Y	E	A	S	V	A	R	219
671	TAC	AGC	GAC	CTG	CCC	GAG	AAG	GCC	AAC	<u>AGA</u>	<u>TCT</u>	<u>CCC</u>	<u>GGG</u>	ATG	GTG	AGC	718
220	Y	S	D	L	P	E	K	A	N	R	S	P	G	M	V	S	235
719	GGC	CTG	CTG	AAG	GAG	AGC	ATG	CGC	ATC	AAG	ATG	TAC	ATG	GAG	GGC	ACC	766
236	G	L	L	K	E	S	M	R	I	K	M	Y	M	E	G	T	251
767	GTG	AAC	GGC	CAC	TAC	TTC	AAG	TGC	GAG	GGC	GAG	GGC	GAC	GGC	AAC	CCC	814
252	V	N	G	H	Y	F	K	C	E	G	E	G	D	G	N	P	267
815	TTC	GCC	GGC	ACC	CAG	AGC	ATG	CGG	ATC	CAC	GTG	ACC	GAG	GGC	GCC	CCC	862
268	F	A	G	T	Q	S	M	R	I	H	V	T	E	G	A	P	283

09976673-104204

102101-102101

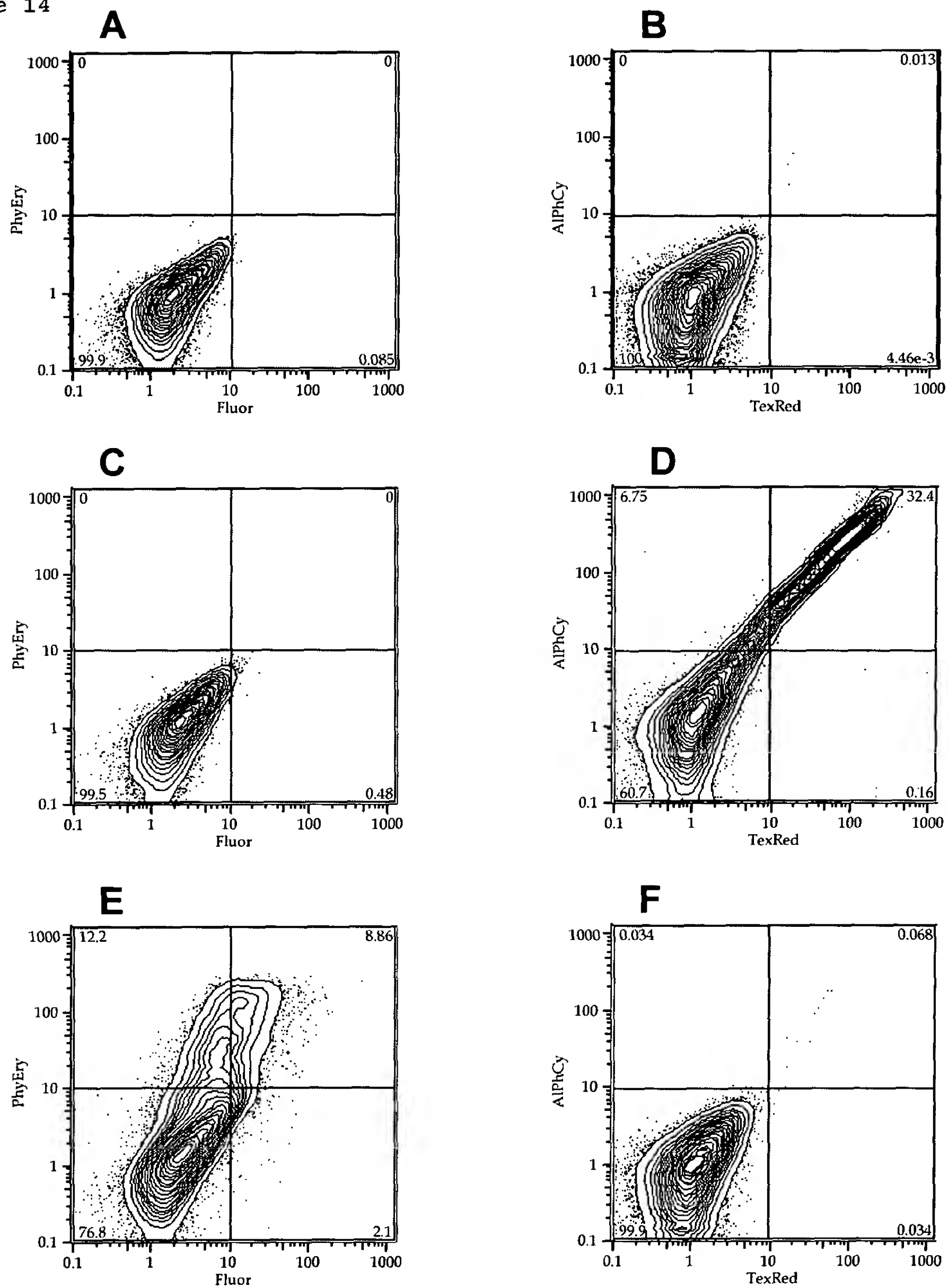
Figure 13 (continued)

863	CTG	CCC	TTC	GCC	TTC	GAC	ATC	CTG	GCC	CCC	TGC	TGC	GAG	TAC	GGC	AGC	910
284	L	P	F	A	F	D	I	L	A	P	C	C	E	Y	G	S	299
911	AGG	ACC	TTC	GTG	CAC	CAC	ACC	GCC	GAG	ATC	CCC	GAC	TTC	TTC	AAG	CAG	958
300	R	T	F	V	H	H	T	A	E	I	P	D	F	F	K	Q	315
959	AGC	TTC	CCC	GAG	GGC	TTC	ACC	TGG	GAG	AGA	ACC	ACC	ACC	TAC	GAG	GAC	1006
316	S	F	P	E	G	F	T	W	E	R	T	T	T	Y	E	D	331
1007	GGC	GGC	ATC	CTG	ACC	GCC	CAC	CAG	GAC	ACC	AGC	CTG	GAG	GGC	AAC	TGC	1054
332	G	G	I	L	T	A	H	Q	D	T	S	L	E	G	N	C	347
1055	CTG	ATC	TAC	AAG	GTG	AAG	GTG	CTG	GGC	ACC	AAC	TTC	CCC	GCC	GAC	GGC	1102
348	L	I	Y	K	V	K	V	L	G	T	N	F	P	A	D	G	363
1103	CCC	GTG	ATG	AAG	AAC	AAG	AGC	GGC	GGC	TGG	GAG	CCC	AGC	ACC	GAG	GTG	1150
364	P	V	M	K	N	K	S	G	G	W	E	P	S	T	E	V	379
1151	GTG	TAC	CCC	GAG	AAC	GGC	GTG	CTG	TGC	GGC	CGG	AAC	GTG	ATG	GCC	CTG	1198
380	V	Y	P	E	N	G	V	L	C	G	R	N	V	M	A	L	395
1199	AAG	GTG	GGC	GAC	CGG	CGG	CTG	ATC	TGC	CAC	CAC	TAC	ACC	AGC	TAC	CGG	1246
396	K	V	G	D	R	R	L	I	C	H	H	Y	T	S	Y	R	411
1247	AGC	AAG	AAG	GCC	GTG	CGG	GCC	CTG	ACC	ATG	CCC	GGC	TTC	CAC	TTC	ACC	1294
412	S	K	K	A	V	R	A	L	T	M	P	G	F	H	F	T	427
1295	GAC	ATC	CGG	CTG	CAG	ATG	CTG	CGG	AAG	GAG	AAG	GAC	GAG	TAC	TTC	GAG	1342
428	D	I	R	L	Q	M	L	R	K	E	K	D	E	Y	F	E	443
1343	CTG	TAC	GAG	GCC	AGC	GTG	GCC	CGG	TAC	AGC	GAC	CTG	CCC	GAG	AAG	GCC	1390
444	L	Y	E	A	S	V	A	R	Y	S	D	L	P	E	K	A	459
1391	AAC	AGA	ACT	CGA	GCT	ATG	GAT	GAT	GAT	ATC	GCC	G...					1424
460	N	R	T	R	A	M	D	D	D	I	A...						470

actin

(SEQ ID NOS. 17 & 18).

Figure 14



099763-10101

Figure 15

Heteractis crispa chromoprotein hcCP mut C148S

C148S according to GFP numbering

C143S according to self-numbering.

ATGGCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACAT
M A G L L K E S M R I K M Y M

GGAAGGCACGGTTAATGGCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATT
E G T V N G H Y F K C E G E G D G N P F

TACAGGTACGCAGAGCATGAGGATTCATGTCACCGAAGGGGCTCCATTACCATTTCCTT
T G T Q S M R I H V T E G A P L P F A F

CGACATTTTGGCACCGTGTGTGAGTACGGCAGCAGGACCTTTGTCCACCATAACGGCAGA
D I L A P C C E Y G S R T F V H H T A E

GATTCCCGATTTCTTCAAGCAGTCTTTCCCTGAAGGCTTTACTTGGGAAAGAACCACAAC
I P D F F K Q S F P E G F T W E R T T T

CTATGAAGATGGAGGCATTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAACTGCCT
Y E D G G I L T A H Q D T S L E G N C L

TATATACAAGGTGAAAGTCCTTGGTACCAATTTTCCTGCTGATGGCCCCGTGATGAAGAA
I Y K V K V L G T N F P A D G P V M K N

CAAATCAGGAGGATGGGAGCCAAGCACTGAGGTGGTTTATCCAGAGAATGGTGTCTCTGTG
K S G G W E P S T E V V Y P E N G V L C

TGGACGTAATGTGATGGCCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCTCTATAC
G R N V M A L K V G D R R L I C H L Y T

TTCTTACAGGTCCAAGAAAGCAGTCCGTGCCTTGACAATGCCAGGATTTTCATTTTACAGA
S Y R S K K A V R A L T M P G F H F T D

CATCCGCCTTCAGATGCCGAGGAAAAAGAAAGACGAGTACTTTGAACTGTACGAAGCATC
I R L Q M P R K K K D E Y F E L Y E A S

TGTGGCTAGGTACAGTGATCTTCCTGAAAAAGCAAATTGA
V A R Y S D L P E K A N *

SEQ ID NO:23 & 24

09976673-101201

Figure 16

Crispa 44-6 mutant possesses six amino acid substitutions vs. wild type:
A2S,T36A,A65E,C143S,L173H,P201L.

09975673-101201

TCTGGTTTGTGAAAGAAAGTATGCGCATCAAGATGTACAT
S G L L K E S M R I K M Y M

GGAAGGCACGGTTAATGGCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAACCCATT
E G T V N G H Y F K C E G E G D G N P F

TGCAGGTACGCAGAGCATGAGGATTCATGTCACCGAAGGGGCTCCATTACCATTTCCTT
A G T Q S M R I H V T E G A P L P F A F

CGACATTTTGGCACCGTGTGTGCGTACGGCAGCAGGACCTTTGTCCACCATACGGCAGA
D I L A P C C A Y G S R T F V H H T A E

GATTCCCGATTTCTTCAAGCAGTCTTTCCTGAAGGCTTTACTTGGGAAAGAACCACAAC
I P D F F K Q S F P E G F T W E R T T T

CTATGAAGATGGAGGCATTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAACTGCCT
Y E D G G I L T A H Q D T S L E G N C L

←

TATATACAAGGTGAAAGTCCTTGGTACCAATTTTCCTGCTGATGGCCCCGTGATGAAGAA
I Y K V K V L G T N F P A D G P V M K N

CAAATCAGGAGGATGGGAGCCAAGCACTGAGGTGGTTTATCCAGAGAATGGTGTCTCTGTG
K S G G W E P S T E V V Y P E N G V L C

TGGACGTAATGTGATGGCCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCACTATAC
G R N V M A L K V G D R R L I C H H Y T

TTCTTACAGGTCCAAGAAAGCAGTCCGTGCCTTGACAATGCCAGGATTTTCATTTTACAGA
S Y R S K K A V R A L T M P G F H F T D

CATCCGCCTTCAGATGCTGAGGAAAAAGAAAGACGAGTACTTTGAACTGTACGAAGCATC
I R L Q M L R K K K D E Y F E L Y E A S

TGTGGCTAGGTACAGTGATCTTCCTGAAAAAGCAAATTGA
V A R Y S D L P E K A N *

SEQ ID NO:25 & 26

Figure 17

Heteractis crispa chromoprotein wild type (base isoform)

10 20 30 40 50 60
5' ACCATTTGCTTTGGTTCCTTGGCAAACGAAAGTTTAGAACGAAAAGTATGACCCAAATTACA
70 80 90 100 110 120
TCTTCCTCCTGGATCCTTACCATGGCTGGTTTGGTTGAAAGAAAGTATGCGCATCAAGATG
M A G L L K E S M R I K M
130 140 150 160 170 180
TACATGGAAGGCACGGTTAATGGCCATTATTTCAAGTGTGAAGGAGAGGGAGACGGCAAC
Y M E G T V N G H Y F K C E G E G D G N
190 200 210 220 230 240
CCATTTACAGGTACGCAGAGCATGAGGATTCATGTCACCGAAGGGGCTCCATTACCATTT
P F T G T Q S M R I H V T E G A P L P F
250 260 270 280 290 300
GCCTTCGACATTTTGGCACCGTGTTGTGAGTACGGCAGCAGGACCTTTGTCCACCATACG
A F D I L A P C C E Y G S R T F V H H T
310 320 330 340 350 360
GCAGAGATTCCCGATTTCTTCAAGCAGTCTTCCCTGAAGGCTTTACTTGGGAAAGAACC
A E I P D F F K Q S F P E G F T W E R T
370 380 390 400 410 420
ACAACCTATGAAGATGGAGGCATTCTTACTGCTCATCAGGACACAAGCCTGGAGGGGAAC
T T Y E D G G I L T A H Q D T S L E G N
430 440 450 460 470 480
TGCCTTATATACAAGGTGAAAGTCCTTGGTACCAATTTTCCTGCTGATGGCCCCGTGATG
C L I Y K V K V L G T N F P A D G P V M
490 500 510 520 530 540
AAGAACAAATCAGGAGGATGGGAGCCATGCACTGAGGTGGTTTATCCAGAGAATGGTGTC
K N K S G G W E P C T E V V Y P E N G V
550 560 570 580 590 600
CTGTGTGGACGTAATGTGATGGCCCTTAAAGTCGGTGATCGTCGTTTGATCTGCCATCTC
L C G R N V M A L K V G D R R L I C H L
610 620 630 640 650 660
TATACTTCTTACAGGTCCAAGAAAGCAGTCCGTGCCTTGACAATGCCAGGATTTTCATTTT
Y T S Y R S K K A V R A L T M P G F H F
670 680 690 700 710 720
ACAGACATCCGCCTTCAGATGCCGAGGAAAACGAAAGACGAGTACTTTGAACTGTACGAA
T D I R L Q M P R K T K D E Y F E L Y E
730 740 750 760 770 780
GCATCTGTGGCTAGGTACAGTGATCTTCCTGAAAAAGCAAATTGATTGTTCCCAGTGACA
A S V A R Y S D L P E K A N *
790 800 810 820 830 840
CCAGACTGCTGTCAGCTTTTGGTTAAAGCCCGAAAGACAAAAGGACATTTGTAGTTTAGT
850 860 870 880 890 900
TTATATTTCCCTTTTCATTTGTGAATCAACATTGTACTCTCTGTAAACCTTTAAAATGCTC
910
CATTAACCT 3' (SEQ ID NOS: 27 & 28)

099363-101201